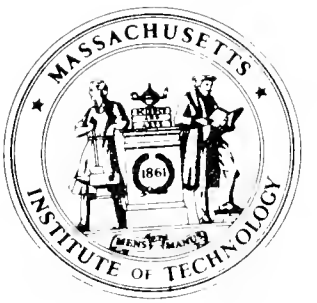


LIBRARY
OF THE
MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

MASS. INST. TECH.
SEP 17 1970
DEWEY LIBRARY

THE STRUCTURE AND DISCRIMINANT POWER OF ATTITUDES

by

Yvan Allaire

Working Paper No. 481-70
June 1, 1970

[illegible]

This paper presents a method of obtaining attitude measures which are relevant to marketing strategy.

The methodology presented here is based on a procedure suggested by Howard and Sheth in their book The Theory of Buyer Behavior (6) and is illustrated with the results of an actual study.

Howard and Sheth's Approach

Briefly stated, Howard and Sheth's procedure runs as follows:

-From the prior beliefs of marketing, personnel and/or small-scale, exploratory, projective surveys, determine the Purchase Motives relevant to the product category under study

-Construct what is essentially a semantic differential but with one bipolar scale for each previously identified Motive. An n x m matrix of raw data is then obtained for each relevant brand (where n = number of respondents; m = number of motives)

-Manipulate this matrix in the principal component sense so as to obtain the underlying structure of the data. The components, or factors, thus identified are labeled "Choice Criteria".

-For each respondent, a vector of "attitude scores" is then computed in a manner very similar to the derivation of factor scores in the traditional factor analysis.

-The attitude scores are then used to measure the Euclidean distance between brands over the set of Choice Criteria.

While this provides a useful framework, the following weaknesses are noteworthy:

(1) Ambiguous Loadings. The rather informally determined Motives are assumed to be perfectly measured by a single bipolar scale. Quite conceivably,

536351

such a scale could result in ambiguous loadings on the Choice Criteria (i.e. substantial loadings on a number of Choice Criteria). Would this indicate that this particular motive is not relevant (which should raise questions about the prior estimate of Purchase Motives), or simply that the wording used to express the motive was ambiguous and did not measure what it was supposed to measure?

(2) No Prior Hypothesis Concerning Choice Criteria. Since prior information is available and since Factor Analysis and related techniques rely to a considerable extent on the subjective judgment of the analyst, a compelling case may be made for the linking of Purchase Motives to Choice Criteria by a prior theory stating (1) the anticipated number of Choice Criteria and (2) the expected pattern of loadings; i.e. which Motives will group together.

The acceptance of whatever results are produced by the manipulation of the matrix of raw data deserves a strong caveat, particularly within the present framework. The prior theory, at least, provides a standard against which the output of the statistical analysis may be measured.

Major differences between prior theory and results have to be investigated and may lead to a new prior theory. Minor divergences may be reconciled by the Bayesian argument that prior information should be updated by sample evidence.

Armstrong and Soelberg (2 and 2A) provide very striking illustrations of the peril from using Factor Analysis without any prior hypothesis (or theory).

(3) Non-Conformities in Loading Pattern. The pattern of loading resulting from the matrix operations may differ from brand to brand. Howard and Sheth argue that if such is the case, then the "various brands are not really elements of the same product class from the buyers' point of view..." (9, p. 213).

However, it must be recognized that there are two ways in which the factor pattern may diverge:

i) The same number of significant factors is obtained but the variables (motives) load differently on the factors so as to make the description of factors (choice criteria) different from one brand to another. In such a case, Howard and Sheth's argument may be acceptable.

ii) The number of significant factors may vary from one brand to another. Therefore, a more plausible argument would be that, despite the fact that the brands are part of a buyer's "evoked set" (i.e. he is aware of the brand and would consider buying it), there is considerable variation in familiarity over all brands considered. The less experience with, and knowledge of, a brand a buyer has, the less discriminating will his judgment tend to be. This tendency could lead to:

(a) less well known brands producing a smaller number of factors (Choice Criteria);

(b) diverging factor patterns if the analysis of some brand were performed only with respondents who have already purchased that brand and, separately, with respondents who have not up to now bought that particular brand, (though it is part of their evoked set). If such is the case, the question arises as to the meaning of the brand Choice Criteria for all customers, which we then know to be a composite of the (different) Choice Criteria of buyers and non-buyers of that brand.

(4) Significance of the Choice Criteria. The attitude scores are used to measure distances between brands but no suggestion is made to test the significance of such differences (e.g. Wilcoxon's matched pairs signed ranks test, or Wilcoxon's "T" test depending on the form of analysis conducted).

However statistical significance is not sufficient. If the Choice Criteria thus identified are valid, we should be able to use them to distinguish between respondents who regularly buy a specific brand and other respondents, on the basis of their attitude scores toward that brand. Furthermore, each Choice Criterion should be assigned a weight reflecting its relative importance in the choice process. Howard and Sheth claim attitude may not be a good predictor of buyer behavior because of the interference of inhibitory factors (e.g. time pressure, lack of availability, financial constraint, or momentary price change). They might argue that the same factors would reduce the discriminant performance of attitudes. Yet, it would be difficult to have much confidence and interest in a set of Choice Criteria which could not correctly assign a substantial percentage of respondents to the brand which they normally use.

Suggested Procedure

The following method for attitude measurement avoids problems 1, 2, and 4. Problem 3 will be considered separately afterward.

-From past experience, prior knowledge and/or preliminary studies, formulate some hypothesis concerning the number and definition of Choice Criteria relevant to a product category.

-Construct a semantic differential with three or four scales to measure each of the hypothesized Choice Criterion. Specify the pattern of loading expected of each variable. These variables may be similar to the motives in Howard and Sheth but there may be more than one bipolar term per motive and furthermore these are now related to Choice Criteria by a prior theory.

-Perform a Factor Analysis (not Principal Component analysis) consistent with the prior hypothesis. Orthogonal rotation of the axes and estimate of communalities by the method of refactorization would appear methodologically desirable.

-Perform statistical and practical tests to determine whether the hypothesis as to the number of Choice Criteria may be retained. Examine the pattern of loadings for its concordance with prior hypotheses.

-If some of the prior hypotheses are rejected, formulate new hypotheses and collect additional data to test them. If this is impractical, the results of the analysis, before used further, should be compared with a "chance" model made up of several randomly generated samples of the same size and statistics as the sample of observed data.

-Compute Attitude scores for each respondent; because of the specific prior (or revised) hypotheses, however, these scores are obtained in a

slightly different manner. The usual "short" regression method is:

$$\underline{Y} = \underline{Z}\,\underline{\beta}$$

where: $\underline{Y}$ is the factor score matrix

$\underline{Z}$ is the matrix of standardized observed scores

$\underline{\beta}$ is the matrix of factor score regression coefficients;

$$\underline{\beta}' = \underline{\alpha}^{-1}\,\underline{S}\,\underline{D}^{-1}$$

where: $\underline{\alpha}'$ is the orthogonal factor loading matrix

$\underline{D}$ is the diagonal matrix of uniqueness; $d_{ii} = 1 - h_i^2$

$$\underline{S} = \underline{I} + \underline{\alpha}'\,\underline{I}^{-1}\,\underline{G}$$

If $\underline{\beta}$ were left unchanged, then each variable would have some effect on all the factors (Choice Criteria); this would introduce unnecessary "noise" since the prior theory specifies which variables define which factors (Choice Criteria). Thus, some elements of $\underline{\beta}$ should be set equal to zero before Attitude scores are computed; i.e. only the regression coefficients which are in agreement with the prior hypothesis are retained.

-Perform a two-way discriminant analysis (preferably on a split sample basis) to estimate the discriminating ability of each Choice Criterion. The discriminant power of the set of Choice Criteria for each brand should be tested on the "fresh" part of the sample or, if impractical to split the sample, against a "chance" model made up of randomly generated samples (3).

The Problem of Divergent Factor Patterns

The problem of non-conformities in the pattern of loadings will be discussed in connection with the following possible outcomes of a Factor Analysis:

Choice Criteria for:	Among:	
Brand X	Regular Buyers of Brand X	(A)
	Non Buyers of Brand X	(B)
Brand Y	Regular Buyers of Brand Y	(C)
	Non Buyers of Brand Y	(D)
Brand Z	Regular Buyers of Brand Z	(E)
	Non Buyers of Brand Z	(F)

Outcome 1: The Choice Criteria are the same in all cases. This is the ideal situation and then, the analysis may be conducted as outlined before.

Outcome 2: A = C ≠ E or some combination of these. Then it may be argued, for example, that Brand Z is not in the same product category as Brands X and Y, from the buyers' point of view. Though this may be a finding of great usefulness, it means that inter-brand comparisons cannot be made with Brand Z.

Outcome 3: A ≠ B, C ≠ D, E ≠ F but A = C = E; i.e. loading patterns differ as between buyers and non-buyers of a brand, presumably as a result of varying degree of familiarity and experience with a given brand among respondents. In such an eventuality, discriminant analysis of Attitude scores toward a given brand to classify respondents into the "buyer" or "non-buyer" group makes little sense. The analysis of Attitude scores could then rely on various Euclidean distance measures: e.g. comparison between scores of Brand X among non-buyers of Brand X, Brand Y among non-buyers of Brand Y, Brand Z among non-buyers of Brand Z.

Illustration of Proposed Methodology

The study was designed to measure attitudes toward a fast growing Savings and Loan institution (with a populist and nationalistic flavor) and a Chartered Bank, among a sample of 111 French Canadian wage-earners residing in Sherbrooke, Province of Quebec, Canada. One of the many instruments used to elicit the attitudes and motivations of respondents was a ten-variable semantic differential*:

1) Moderne..... Vieux jeu
2) Rapide (service)..... Lente (service)
3) Professionnelle..... Amateur
4) Mauvaise..... Bonne
5) Active..... Passive
6) Favorise les gros..... Favorise les petits
7) Faible (financièrement)..... Forte (financièrement)
8) Plaisante..... Déplaisante
9) Aide les Canadiens Francais..... Nuit aux Canadiens Francais
10) Petite..... Grosse

Two points must be made at this juncture.

(a) It may be argued that the choice process examined here is very different from the product buying behavior studied by Howard and Sheth. While conceding that some adjustments may be necessary, it seems to us that a theory of buyer behavior which could not accommodate these variations would

---

*A tame English translation of the variables would be: (1) Modern - old-fashioned; (2) Fast (service) - Slow (service); (3) Professional - Amateur; (4) Bad - Good; (5) Active - Passive; (6) Favors the wealthy - Favors the wage-earners (?); (7) Weak (financially) - Strong; (8) Pleasant - Unpleasant; (9) Helps French Canadians - Harmful (?) to French Canadians; (10) Small - Big.

be unduly fragile and restricted.

(b) The study reported here was carried on during the fall of 1967. Therefore, it does not represent an attempt to replicate (with improvements) the procedure suggested by Howard and Sheth. The flexibility of our analysis is thus limited by the data collection process used at the time.

The following hypotheses were formulated:

H_1: The information contained in the ten variables is adequately summarized by three factors.

H_2: These factors may be labeled as follows:

- an evaluative factor measured by the variables 1, 3, 4, 8,
- a dynamism factor measured by variables 2, 5, 7, 10,
- an affiliation factor, indicating the extent to which the institution was perceived as supportive and "close" to French Canadian wage-earners; measured by variables 6 and 9.

It might have been advisable to include additional Choice Criteria such as convenience of location and ease of borrowing. However this study was concerned with identifying, if present, more intangible considerations.

A Factor Analysis of the Pearson product-moment correlation matrix was performed using the Principal-factor method and orthogonal rotation of the axes under Kaiser's Varimax criterion. The communalities were estimated by using the square of the multiple correlation coefficient as the original estimate and iterating until communalities converged to two decimal points.

The Results of the Factor Analysis

Tables 1 and 2 present the results of this statistical treatment. Both the formal statistical test (Harman 1960, p. 254) and the often more

relevant practical considerations* militate in favor of the acceptance of H_1; there seem to be three factors underlying the raw data.

A close scrutiny of the pattern of loadings forces a rejection of H_2 in favor of a slightly amended formulation:

-Factor 1 measures a dynamic dimension; i.e. the extent to which the institution is perceived as modern, pleasant, providing quick service (variables 1, 2, 5, 8).

-Factor 2 seems to measure an element of potency, or financial soundness and stability, with high loadings on variables 2, 7 and 10.

-Factor 3, as per prior hypothesis, could be albeled an affiliation dimension with high loadings on variables 6 and 9.

The results are very similar for the two institutions with the exception of variable 3 which fluctuates between Factors 1 and 2.

The study could end at this point with the recommendation that a new set of data be collected to infirm of confirm the Hypothesis 2 as reformulated above. Such a "pure" methodological approach may conflict with cost and time considerations, particularly in this case where results and prior hypotheses do not differ greatly. However, if the analysis is to be pursued with results which do not entirely agree with our prior hypotheses, there must be some insurance against the risk of using results which are merely a statistical aberration. The following procedure should meet that requirement.

Simulated Samples

Normally distributed random numbers were generated fo form ten samples of the same size, mean and variance as the actual data for the Savings and Loan

*See Appendix B.

Table I

Factor Pattern - Savings and Loan Institution

	Common Factors			Communalities
	F_1	F_2	F_3	h_i^2
Variable 1	.745	.370	.012	.692
2	.630	.227	.002	.449
3	.493	.608	.087	.621
4	.318	.440	.165	.322
5	.716	.282	.134	.610
6	.039	.124	.574	.347
7	.330	.587	.120	.468
8	.743	.297	.230	.693
9	.222	.048	.717	.565
10	.187	.654	.074	.467
Contribution of Factor i(V_i)	2.58	1.62	.92	$\sum_{i=1}^{10} h_i^2 = 5.23$
% of common variance accounted by F_i; $100(V_i/5.23)$	49.3%	31%	17.6%	

Table 2

Factor Pattern - Bank X

	Common Factors			Communalities
	F_1	F_2	F_3	h_i^2
Variable 1	.595	.104	.207	.406
2	.646	.257	.212	.528
3	.502	.351	.035	.377
4	.394	.309	.378	.394
5	.728	.046	.135	.550
6	.449	.070	.430	.212
7	.172	.777	.073	.639
8	.628	.233	.220	.496
9	.112	.105	.614	.400
10	.183	.696	.011	.517
Contribution of Factor i (V_i)	2.44	1.52	.83	$\Sigma h_i^2 = 4.52$
% of common variance accounted by F_i; $100(V_i/4.52)$	54%	33.6%	18.2%	

Table 3

Factor Analysis Results
"Best" Simulated Sample

	F_1	F_2	F_3	h_i^2
Variable 1	.29	.07	.12	.10
2	.14	.14	.21	.08
3	.09	.31	.02	.11
4	.12	.01	.41	.18
5	.43	.15	.05	.21
6	.13	.01	.08	.02
7	.31	.05	.05	.10
8	.01	.43	.11	.20
9	.05	.28	.32	.19
10	.07	.05	.21	.05

Table 4

Distribution of Factor Loadings

	For 10 Simulated "Samples"	For Actual Sample
No. of loadings between .80 - 1.00	0	0
.70 - .79	0	4
.60 - .69	0	3
.50 - .59	2	2
.40 - .49	18	2

Loadings below .40 were never considered significant in our analysis.

institution. These samples were factor analyzed using the method and criteria outlined before. Table 3 presents the results for the simulated sample showing the pattern most closely approximating Table 1. Table 4 compares the factor loadings. Both tables strongly support the claim that the results obtained with the actual data are probably significant.

Discriminant Power of Attitudes

Attitude scores were computed for each respondent as per the procedure outlined before:

$$\underline{Y} = \underline{Z}\ \underline{B}$$

For the Savings and Loan institution, for example, $\underline{B}$ is:

	F_1	F_2	F_3
1	**.337**	.001	.150
2	**.189**	.056	.074
3	.005	**.348**	.036
4	.023	.146	.044
5	**.275**	.081	.008
6	.094	.052	**.321**
7	.070	**.287**	.014
8	**.344**	.111	.120
9	.013	.108	**.578**
10	.151	**.386**	.008

In accordance with our revised hypothesis, the coefficients which are not boxed were set equal to zero before attitude scores were computed.

A discriminant analysis of these scores was performed to classify respondents as customers of the S. & L. institution or bank on the basis of their attitudes toward each institution. Tables 4 and 5 present there results. The standardized coefficients of the discriminant functions indicate the discriminating ability of each factor (choice criterion) and thus its importance in the composition of attitudes toward each institution. The percentage of correctly classified respondents should be compared with the following naive model (9):

Naive Model

Let the prior probability of assigning an observation to the S. & L. group be denoted by $P(S.L.)$ and to the bank group by $P(B)$. Let the probability of assigning an observation correctly to the group to which it belongs be denoted by $P(K)$. Then

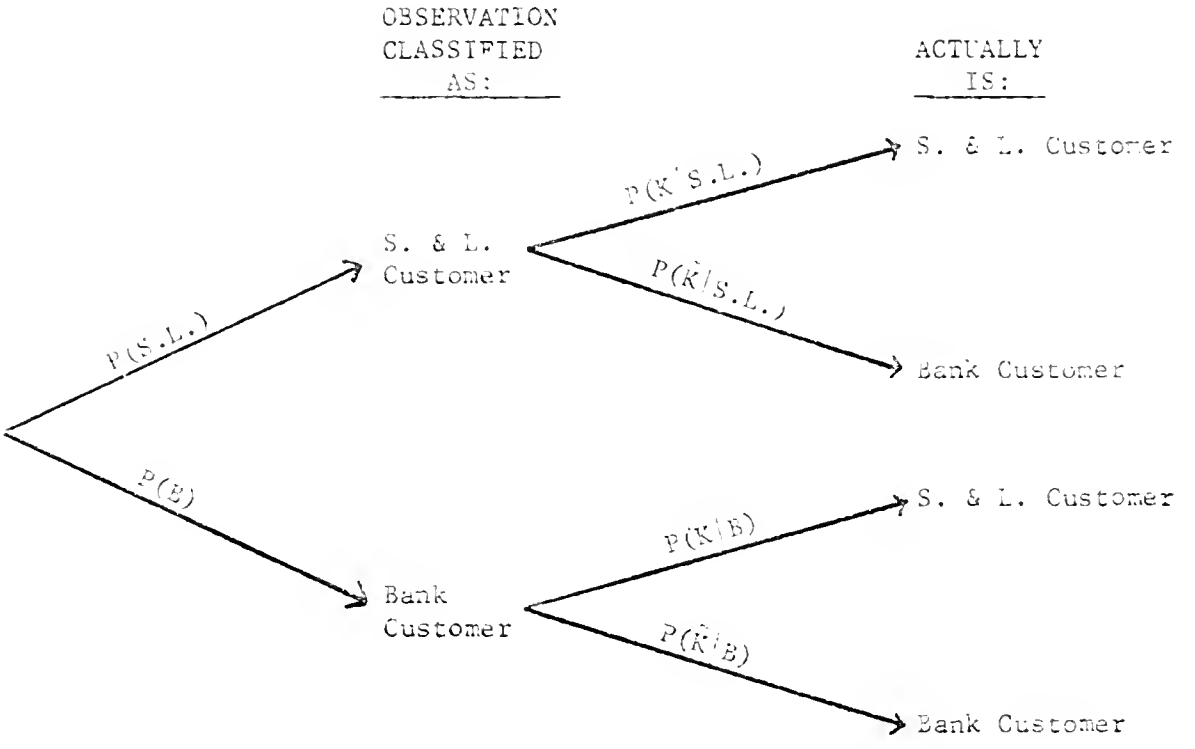


Table 4

Discriminant Analysis of Attitudes Toward
The Savings and Loan Institution

Discriminant Function (standardized)*:

.057 Y_1 (Dynamism) + .334 Y_2 (Potency) -- .731 Y_3 (Affiliation)

Confusion Table:

	Classified as: S. & L. Customers	Bank Customers	TOTAL
S. & L. Customers	65	6	71
Bank Customers	18	22	40
			111

Percent correctly classified 87/111 = 78.4%

Table 5

Discriminant Analysis of Attitudes Toward Bank X

Discriminant Function (standardized):

.606 Y_1 (Dynamism) + .722 Y_2 (Potency) + .180 Y_3 (Affiliation)

Confusion Table:

	Classified As: S. & L. Customers	Bank Customers	TOTAL
S. & L. Customers	60	11	71
Bank Customers	13	27	40
			111

Percent correctly classified 87/111 = 78.4%

On an "a priori" basis, $P(S.L.) = 71/111$,

$P(B) = 40/111$,

$P(K|S.L.) = 71/111$,

$P(K|B) = 40/111$.

Then $P(K) = P(S.L.) \cdot P(K|S.L.) + P(B) \cdot P(K|B)$

.54

Thus, with this naive model, we would expect to correctly classify only 54% of the observations.

However, since the results of tables 4 and 5 are obtained with the same data from which the discriminant functions were computed, they probably overstate to some extent the discriminant power of the variables. Nevertheless, even with some important Choice Criteria omitted (convenience, interest rates, ease of borrowing), the Attitude scores appear somewhat better than a naive model at identifying customers. These results should be tested with a "fresh" sample or, at least, against a "chance" model made up of simulated results.

Interpretation of Discriminant Functions

Valuable information may be surmised from the above results:

--While both groups (S. & L. customers and bank customers) clearly favor the institution they patronize on the "Dynamism" dimension, the attitude structure is much more complex for the other two dimensions.

--On the "Potency" factor, both groups were in substantial agreement when evaluating the S. and L. but disagreed substantially when evaluating Bank X. The conclusion drawn (and supported by further analysis of the data) is that while S. and L. customers perceived both institutions as similarly "potent", the bank customers viewed the bank as much more powerful and strong than the S. and L., yet perceiving the latter at the same level of potency as did the S. and L. customers.

--On the "Affiliation" dimension, a similar comment may be made. Both groups perceive Bank X rather similarly on that factor but the S. and L. group perceived their institution as quite different from the Bank on that dimension. The bank group perceives both institutions as quite similar.

The managerial implications of the above comments are fairly obvious and important.

Conclusion

Using a procedure suggested by Howard and Sheth as a reference, a methodology to measure the structure and discriminant power of attitudes was outlined and illustrated with an actual study of attitudes toward some financial institutions.

Appendix A

Factor Analysis: The Statistical Model

A Factor Analysis of the Pearson product-moment correlation matrix was performed using the Principal-factor method and orthogonal rotation under Kaiser's Varimax criterion.* Under hypothesis H_1, the following data-generating process is assumed:

(1) $$\underline{X} = \underline{A}\,\underline{Y} + \underline{U}$$

where $\underline{X}$ = vector consisting of "p" observed responses

$(X_1, X_2, \ldots, X_p)$; here $p = 10$

$\underline{Y}$ = vector of (nonobservable) common factor variates

$(Y_1, Y_2, \ldots, Y_m)$; here, because of H_1, $m = 3$.

$\underline{A}$ = matrix of factor loadings ($p \times m$) which reflect the importance of the jth factor in the composition of the ith response. (a_{ij})

$\underline{U}$ = vector of (nonobservable) specific factors and errors

$(u_1, u_2, \ldots, u_p)$

It is assumed that:

-the variates in $\underline{Y}$ are independent and normally distributed with 0 mean and unit variance

-the elements of $\underline{U}$ are independent and normally distributed with 0 mean and variance [var $(u_i) = \psi_i$] or in matrix terms

---

*Oblique rotation using the Promax method was also performed. However, since the resulting pattern matrix was only slightly "cleaner", only the orthogonal solution was retained.

$$\underline{\Psi} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & \psi_2 & & \\ \vdots & & & \\ 0 & & & \psi_{10} \end{bmatrix}$$

Therefore the model states that any response is a composite of the effects of the three common factors and a unique element:

$$(2) \qquad \begin{aligned} X_1 &= a_{1,1}Y_1 + a_{1,2}Y_2 + a_{1,3}Y_3 + v_1 \\ &\vdots \qquad\quad \vdots \qquad\quad \vdots \qquad\quad \vdots \qquad\quad \vdots \\ X_{10} &= a_{10,1}Y_1 + a_{10,2}Y_2 + a_{10,3}Y_3 + v_{10} \end{aligned}$$

From the above assumptions, it is clear that

$$(3) \qquad \text{Var}(X_i) = \sigma_i^2 = a_{i1}^2 + a_{i2}^2 + a_{i3}^2 + \psi_i$$

$$(4) \qquad \text{Covariance}(X_{ij}) = \sigma_{ij} = a_{i1}a_{j1} + \ldots + a_{i3}a_{j3}$$

In matrix form, we have:

$$(5) \qquad \underline{\Sigma} = \underline{A}\,\underline{A}' + \underline{\Psi}$$

If the correlation matrix is used as input, then

$$(6) \qquad \underline{R} = \underline{A}\,\underline{A}' + \underline{\Psi}$$

where $r_{ii} = 1.0$ is the equivalent of σ_i^2

and r_{ij} is related to σ_{ij}.

Thus the diagonal of the correlation matrix is the sum of two components:

1- the communality: $h_i^2 = \sum_{j=1}^{3} a_{ij}^2$

2- the uniqueness: $u_{ii} = 1.0 - h_i^2$

The Principal-Factor solution requires that an estimate of the communalities be provided; i.e. that the value of 1.0 in the diagonal of the correlation matrix be replaced by some number ($r_{ii} < 1.0$). The following procedure for obtaining such estimates, though onerous, is methodologically sound:

1- a first estimate of the communalities is made using the square of the multiple correlation coefficient between the ith variable and the p-1 other variables, since this is the lower bound on the communality (Harman, 1960).

2- a solution of the Factor Analytic model is then obtained which has the following form:

	F_1	F_2	F_3
Variable 1	$a_{1,1}$	$a_{1,2}$	$a_{1,3}$
2	$a_{2,1}$	$a_{2,2}$	$a_{2,3}$
⋮	⋮	⋮	⋮
10	$a_{10,1}$	$a_{10,2}$	$a_{10,3}$

3- a new estimate of communalities is obtained by using $r_{ii} = \sum_{j=1}^{3} a_{i,j}^2$. Go back to step 2.

This process of iteration by refactoring was pursued until convergence, to two decimal points, was achieved in the estimate of communalities.

Appendix B

Goodness of Fit of Model: Statistical Test and Practical Considerations

Statistically we should be concerned with the extent to which the three factors succeed in reproducing the matrix of observed correlations, where r'_{ij}, the reproduced correlation between the ith and j^{th} variable, is given by:

$$(1) \qquad r'_{ij} = \sum_{k=1}^{3} a_{ik} \cdot a_{jk}$$

The following tables present both the observed correlations r_{ij} and the "reproduced" correlations (see tables B1 and B2).

The following (large sample) statistical test may be applied to verify the goodness of fit of the model:

$$(2) \qquad G_m = (N-1) \ln \frac{|\Lambda\Lambda' + \Psi|}{|\Lambda|} \qquad (5, p. 380)$$

where G_m is asymptotically distributed as χ^2 with degrees of freedom $\nu = \frac{1}{2}[(p-m)^2 + p - m]$.

The hypothesis of the common factors is rejected if $G_m > \chi^2_{\alpha,\nu}$, and accepted otherwise. Expression (2) is well approximated by computing:

$$(3) \qquad G_m = (N-1) \sum_{i=1}^{10} \sum_{j=i}^{10} \frac{(r_{ij} - r'_{ij})^2}{(1-h_i^2)(1-h_j^2)}$$

The results, shown at the bottom of tables B1 and B2, indicate that there is some statistical justification in retaining hypothesis 1.*

---

*A further statistical consideration: for the case of Principal Components Analysis (i.e. (roughly) when the value of 1.0 is left in the diagonal of the correlation matrix), Guttman's (4) lower bound theorem shows that eigenvalues with roots less than 1.0 are statistically insignificant. A P.C. analysis of our data produced only three components with roots larger than 1.0.

Table B1

Savings and Loan Institution--Observed (r_{ij}) and Reproduced (r'_{ij}) Correlations

	r'_{ij}									
	1	2	3	4	5	6	7	8	9	10
Var 1	--	.63	.59	.31	.62	.07	.49	.66	.18	.40
2	.55	--	.46	.27	.50	.05	.29	.56	.14	.29
3	.59	.45	--	.45	.51	.05	.58	.55	.25	.47
4	.40	.30	.44	--	.39	.14	.53	.53	.17	.41
5	.64	.52	.54	.37	--	.06	.46	.67	.27	.31
6	.02	.01	.11	.14	.08	--	.12	.09	.43	.15
7	.47	.34	.53	.39	.42	.13	--	.40	.21	.44
8	.67	.34	.57	.41	.65	.14	.45	--	.36	.35
9	.19	.15	.20	.21	.27	.41	.19	.34	--	.09
10	.38	.27	.50	.36	.33	.12	.45	.35	.13	--

$G_3 = 36.30$ $\nu = 28$ deg. of freedom

$X^2_{.10,28d.f.} = 37.9$

Table b2

Bank X--Observed (r_{ij}) and Reproduced (r'_{ij}) Correlations

Var.	r'_{ij}	r_{ij} 1	2	3	4	5	6	7	8	9	10
	1	--	.43	.44	.33	.45	.14	.20	.43	.25	.14
	2	.46	--	.34	.34	.50	.27	.31	.63	.19	.37
	3	.34	.42	--	.37	.41	.06	.37	.33	.10	.35
	4	.35	.41	.32	--	.44	.20	.36	.56	.33	.27
	5	.47	.51	.39	.35	--	.13	.17	.48	.15	.13
	6	.17	.17	.07	.20	.16	--	-.06	.14	.27	.03
	7	.20	.33	.36	.33	.17	.01	--	.32	.18	.56
	8	.44	.51	.40	.40	.50	.17	.30	--	.25	.27
	9	.21	.23	.11	.31	.17	.27	.15	.23	--	.05
	10	.18	.29	.34	.28	.16	-.03	.57	.27	.09	--

$G_3 = 34.87$ $\nu = 28$ degrees of freedom

$X^2_{.10,28 d.f.} = 37.9$

Practical considerations could often outweigh the statistical evidence, particularly if additional factors have no theoretical, empirical or logical meaning and their contribution to the total explained variance is small. When a four-factor model was adjusted to the data, the fourth factor contained only small loadings and proved to be very ambiguous. Thus, the practical and statistical considerations both lead to a fairly confident acceptance of hypothesis 1.

Appendix C

Comparison of simulated "samples" with actual sample

A rough index of congruence between the factor loadings of the simulated "samples" and the actual sample may be:

$$(1) \qquad C_k = \sum_{i=1}^{10} \sum_{j=1}^{3} (a'_{ij} - a_{ij})^2$$

where C_k = index of congruence between simulated sample k (k=1,2,...,10) and actual sample

a'_{ij} = factor loading of variable i on factor j for the simulated sample

a_{ij} = factor loading of actual sample.

The values of C_k were:

k =	1	2	3	4	5	6	7	8	9	10
C_k	5.65	7.81	6.98	5.68	6.75	6.52	5.45	7.53	5.76	5.39

"Sample" 10 is the most congruent of the samples and is presented in table 3.

Appendix D

Discriminant analysis: the statistical model

Hypothesis: respondents may be classified into their respective groups (i.e. customers or non-customers) on the basis of their attitude factor scores.

Discriminant analysis was used to test this hypothesis.

If it is assumed that:

(1) the observations are drawn from a multivariate normal distribution

(2) with different (but unknown) mean vectors $(\mu_1^{(1)}, \mu_2^{(1)}, \mu_3^{(1)})$ and $(\mu_1^{(2)}, \mu_2^{(2)}, \mu_3^{(2)})$

(3) and common variance-covariance matrix (also unknown):

$$\Sigma_1 = \Sigma_2$$

then the following expression is obtained (1):

$$C = \underline{x}'S^{-1}(\bar{\underline{x}}^{(1)} - \bar{\underline{x}}^{(2)}) - \tfrac{1}{2}(\bar{\underline{x}}^{(1)} + \bar{\underline{x}}^{(2)})'S^{-1}(\bar{\underline{x}}^{(1)} - \bar{\underline{x}}^{(2)}) \qquad (1)$$

where $\underline{S}$ = sample variance-covariance matrix (symmetric) computed from (2)

$$\underline{S} = \frac{\sum_{j=1}^{3}\sum_{k=j}^{3}\sum_{i=1}^{N_1+N_2}(y_{ij} - \bar{\bar{y}}_j)(y_{ik} - \bar{\bar{y}}_k)}{N_1 + N_2 - 2} \qquad (2)$$

$\bar{\bar{y}}_j$ = mean of variable j ($j=1,2,3$) for all observations

N_1 = no. of observations in group 1 (71)

N_2 = no. of observations in group 2 (40)

The first part of expression (1) is the discriminant function.

However, these three assumptions should be examined.

[illegible] ition, [illegible] normal [illegible].

Assumption (2) [illegible] the following null hypothesis must be rejected:

$$H_0: \underline{\ }^{(1)} = \underline{\ }^{(2)}$$

The generalized Mahanalobis D^2 statistic is convenient to test the null hypothesis:

$$D^2 = \sum_{j=1}^{3} \sum_{k=1}^{3} \underline{S}^{-1} \sum_{g=1}^{2} \ldots (\bar{y}_{jg} - \bar{y}_j)(\bar{y}_{kg} - \bar{y}_k)$$

This statistic can be used as a chi-square distributed variate with $m(g-1)$ degrees of freedom, where g = no. of groups. If $D^2 > \chi^2_{a,\ldots}$ H_0 is rejected.

Assumption (3) implies acceptance of the null hypothesis:

$$H_0: \underline{\ }^{(1)} = \underline{\ }^{(2)}$$

against the alternative:

$$H_1: \underline{\ }^{(1)} \neq \underline{\ }^{(2)}$$

The following statistic (8) is recommended:

(17) $$M = \sum_{i=1}^{8} n_i \ln |\underline{S}| - \sum_{i=1}^{8} n_i \ln |\underline{S}_i|$$

where: $n_i = N_i - 1$

S_i - determinant of variance-covariance matrix of entire sample

Σ_i determinant of variance-covariance matrix for group i (i=1,....

$$C^{-1} = 1 - \frac{[illegible]}{[illegible](g-1)} \left(\sum_{i=1}^{[illegible]} \frac{[illegible]}{n_i} - \frac{[illegible]}{\sum_{i=1}^{[illegible]} n_i} \right)$$

where: p = no. of variables (= 5)

g = no. of groups (= 2)

Then MC^{-1} is approximately distributed as a χ^2 variate with degrees of freedom, $\nu = \frac{1}{2}(g-1)p(p+1)$. H_0 may be retained if $MC^{-1} \leq \chi^2_{\alpha,\nu}$. "The approximation appears to be good if "g" and "p" do not exceed 4 or 5, and each n_i is perhaps 20 or more" ([illegible], p. [illegible]).

Table D1 summarizes these statistical results. We find that both assumptions (2) and (3) are quite tenuous. The differences between mean vectors is not highly significant (level of confidence 10%) and the hypothesis of common covariance is all but rejected. However discriminant analysis has been found to be quite robust to breach of the latter assumption (8).

Table D1

Discriminant Analysis on [illegible] Scores Toward.

	Savings and Loan Institution	Banks
Generalized Mahalanobis D^2 =	7.29	6.[illegible]
$\chi^2_{.10,3}$ = 6.25		
C^{-1} =	17.[illegible]	19.9
$\chi^2_{[illegible]}$ = [illegible]		

Appendix E

Discriminant Analysis Classification of Observations

The following Bayesian classification scheme was adopted (Anderson, 1958):

-Compute:

$$C = \underline{Y}' S^{-1}(\bar{\underline{Y}}^{(1)} - \bar{\underline{Y}}^{(2)}) - \tfrac{1}{2}(\bar{\underline{Y}}^{(1)} - \bar{\underline{Y}}^{(2)})' S^{-1}(\bar{\underline{Y}}^{(1)} - \bar{\underline{Y}}^{(2)})$$

-If $C > \ln \frac{(1-h)\lambda(1;2)}{h\lambda(2;1)}$

classify observation as coming from the S. &. L. group; otherwise assign it to the bank group.

where: h = prior probability that an observation is from the S. &. L. group. In the present case, h = 71/111 = .64 (which incidentially was also approximately the market share of the S. &. L. in the population).

$\lambda(i,j)$ = loss from classifying an observation from the jth population as one of the ith population. In the present case, $\lambda(1,2)$ was assumed equal to $\lambda(2;1)$.

So if $C > \ln(.5625) = -.57536$, the observation is assumed to come from the S. &. L. group.

REFERENCES

(1) Anderson, T.W., An Introduction to Multivariate Statistical Analysis (New York: John Wiley and Sons, Inc.), 1958.

(2) Armstrong, J.S. and L.O. Soelberg, "A Note On The Interpretation of Factor Analysis", Sloan School of Management, M.I.T., Working Paper no. 254-67.

(3) Armstrong, J.S., "Derivation of Theory by Means of Factor Analysis or Tom Swift and His Electric Factor Analysis Machine", Sloan School of Management, M.I.T., Working Paper No. 265-67.

(4) Duhamel, William F., The Use of Variable Markov Processes as a Basis for The Determination and Analysis of Market Segments, Unpublished Ph.D. dissertation, Stanford University, 1966. (University Microfilms, no. 66-14, 655).

(5) Guttmann, L., "Some Necessary Conditions for Common Factor Analysis", Psychometrika, 1954, 19, 149-161.

(6) Harman, Harry H., Modern Factor Analysis (Chicago: The University of Chicago Press), 1960.

(7) Howard, John A. and J. N. Sheth, The Theory of Buyer Behavior (New York: John Wiley and Sons, Inc.), 1969.

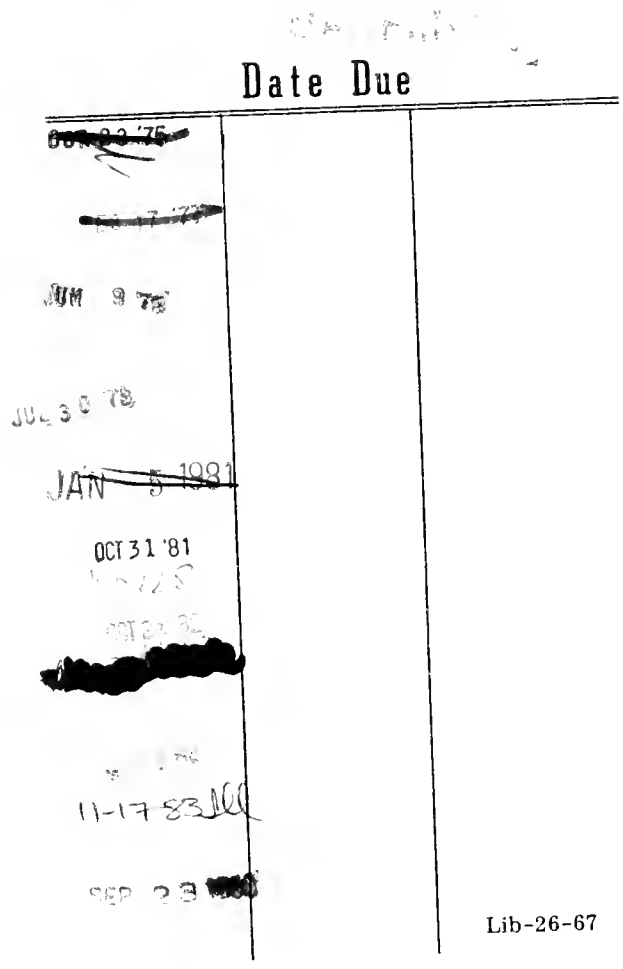
(8) Kendall, M.G., A Course in Multivariate Statistical Analysis, (London: Charles Griffin and Co. Ltd.), 1957.

(9) Morrison, Donald F., Multivariate Statistical Methods, (New York: McGraw-Hill), 1967.

(10) Morrison, Donald F., "On the Interpretation of Discriminant Analysis" Journal of Marketing Research, May 1969, p. 158.

JUL 11 '72

APR 2 '73



Lib-26-67

MIT LIBRARIES DUPL
3 9080 003 905 962
477-7
MIT LIBRARIES DUPL
3 9080 003 906 085
478-70
HD2
.M414
477-70
MIT LIBRARIES DUPL
3 9080 003 875 108
479-70
MIT LIBRARIES DUPL
3 9080 003 875 041
480-70
MIT LIBRARIES DUPL
3 9080 003 905 988
481-70
MIT LIBRARIES DUPL
3 9080 003 906 895
MIT LIBRARIES DUPL
3 9080 003 906 838
MIT LIBRARIES DUPL
9080 003 906 812
MIT LIBRARIES DUPL
3 9080 003 906 770
MIT LIBRARIES DUPL
3 9080 003 826 929
MIT LIBRARIES DUPL
3 9080 003 894 588